

Владимир Брюков

КАК ПРЕДСКАЗАТЬ КУРС ДОЛЛАРА

*Эффективные
методы*

прогнозирования

с использованием

Excel и EViews



Владимир Георгиевич Брюков

Как предсказать курс доллара. Эффективные методы прогнозирования с использованием Excel и EViews

http://www.litres.ru/pages/biblio_book/?art=26105358

Аннотация

Каким будет курс американского доллара завтра, послезавтра, через неделю или через месяц? Сколько будут стоить в ближайшем будущем евро, британский фунт, швейцарский франк и другие валюты? Каких рекомендуемых курсов покупки или продажи при этом следует придерживаться? Это второе дополненное издание книги для тех, кто не только хочет зарабатывать на валютном рынке, но и свести к минимуму свои потери от торговли на валютном рынке. А для этого трейдер должен научиться прогнозировать.

Содержание

Предисловие ко второму дополненному изданию	4
Предисловие к первому изданию	5
Об авторе	11
1.1. Краткая характеристика стационарных и нестационарных случайных процессов	15
1.2. Распознавание стационарности временного ряда с помощью построения его графика	20
Контрольные вопросы и задания к главе 1	36
2.1. Характеристика метода наименьших квадратов и его применение при прогнозировании курса доллара	37
2.2. Решение уравнения регрессии в Excel с учетом фактора времени. Интерпретация и оценка значимости полученных параметров	43
Конец ознакомительного фрагмента.	53

Предисловие ко второму дополненному изданию

В 2011 году издательство КНОРУС и Центр Исследований Платежных Систем и Расчетов выпустили книгу Брюков В. Г. «Как предсказать курс доллара. Эффективные методы прогнозирования с использованием Excel и EViews». К сожалению, далеко не все читатели – в том числе и из-за относительно высокой стоимости печатной версии этой книги – сумели ее приобрести, о чем говорят их письма. В связи с этим автор этих строк решил издать эту книгу вторично в электронном виде, включив во второе издание дополнительный параграф о структурных изменениях в динамике курса доллара к рублю, произошедших в 2014-2015 годах. Кроме того, кто хочет подробнее ознакомиться с моими последними работами по этому поводу, может приобрести недавно вышедшую еще одну мою книгу – с «Как предсказать курс доллара. Расчеты в Excel для снижения риска проигрыша».

Владимир Георгиевич Брюков, независимый аналитик

Предисловие к первому изданию

Каким будет курс американского доллара завтра, послезавтра, через неделю или через месяц? Сколько будут стоить в ближайшем будущем единая европейская валюта, японская йена, британский фунт, швейцарский франк, канадский, австралийский или новозеландский доллар и прочие валюты? Есть ли смысл сегодня (завтра, послезавтра или же через месяц) вкладывать российские рубли в доллары США, евро, фунт, или в йену, либо, напротив, от иностранной валюты нужно поскорее избавляться? Каких рекомендуемых курсов покупки или продажи при этом следует придерживаться?..

Мы живем в такое время, когда ответы на эти животрепещущие вопросы ежедневно приходится искать миллионам рядовых инвесторов и профессиональных трейдеров во всем мире, в том числе и в России. И это вполне понятно, поскольку от правильных ответов на эти актуальные вопросы зависит их будущее благосостояние.

По нашему мнению, и хочется надеяться, что читатель с этим согласится: тот, кто возьмет на себя труд внимательно прочитать, а еще лучше – проштудировать эту книгу – сможет давать более компетентные ответы на эти животрепещущие вопросы.

Следовательно, наша книга предназначена для всех тех,

кто интересуется валютным рынком, кто собирается зарабатывать или уже зарабатывает на этом рынке, и для всех тех, кто хочет научиться делать прогнозы по курсам валют. Впрочем, эта книга будет полезна и интересна не только для валютных инвесторов и трейдеров, но и для студентов, будущая профессия которых, так или иначе, связана с работой в банке, финансовой компании или с операциями на финансовых и товарных рынках. Причем, полученные в этой книге знания пригодятся не только для работы на валютном рынке, но и для работы на других товарных и финансовых рынках, поскольку методика прогнозирования курса валют принципиально не отличается от прогнозирования цен, например, на такие товары как нефть или золото.

В книге детально излагается методики построения стационарных и нестационарных статистических моделей по прогнозированию курса доллара США с использованием программ EViews и Excel. При этом особенно большое внимание уделяется разработке моделей прогнозирования по курсу доллара США с упреждением в один месяц. После освоения этого материала в заключительной главе 7 даны статистические модели по прогнозированию курса доллара с упреждением в две недели и упреждением в одну неделю, а также по прогнозированию курса евро к доллару с упреждением в один день. В этой же заключительной главе рассказывается о методике составления на основе разработанных статистических моделей рекомендуемых цен покупки и про-

дажи валют. Причем, эффективность этих цен рассматривается на конкретных примерах с учетом последних рыночных данных.

Очевидно, что помимо прогнозов по американской валюте и единой европейской валюте многие хотели бы научиться делать прогнозы и по курсам других валют, например, по курсу йены, фунту стерлингов, австралийскому доллару... И этот список интересных для инвестора валют можно расширить, по меньшей мере, на еще два десятка денежных единиц.

В связи с чем, необходимо заметить: внимательно прочитав и освоив материал нашей книги, читатель в дальнейшем может справиться с этой задачей вполне самостоятельно. Поскольку, в основе составления прогнозов по курсам различных валют лежит одна и та же методика, основанная на использовании феноменальных технических возможностей таких мощных вычислительных программ как EViews и Excel.

Материал в нашей книге, состоящей из семи глав, излагается следующим образом. Во-первых, в каждом параграфе главы сначала дается конкретная проблема, связанная прогнозированием курсов валют. Во-вторых, предлагается определенный алгоритм действий с использованием EViews или Excel. В-третьих, даются некоторые математические подробности, разъясняющие суть данного алгоритма действий. В-четвертых, на конкретном примере показывается, как нужно пользоваться данным алгоритмом действий

для решения той или иной проблемы, связанной с прогнозированием курса доллара. И, наконец, в-пятых, в конце каждой главы читателю задаются вопросы, помогающие закрепить пройденный материал.

Очевидно, что после того как наш читатель убедится, что он может заниматься прогнозированием самостоятельно, то ничто не мешает ему воспользоваться EViews или Excel для того чтобы составить прогнозы по курсу, например, японской йены или британского фунта, а также по поводу будущей цены на нефть, золото и прочих товаров.

Вполне понятно, что в эпоху бурного роста электронно-вычислительной техники и всеобщей автоматизации любой желающий овладеть азами валютного прогнозирования может существенно сэкономить время на расчетах, если он сумеет овладеть компьютерными способами обработки статистических данных. Поэтому необходимым условием для нашей совместной работы является умение читателя работать с компьютером на уровне рядового пользователя, а также наличие определенных навыков работы с программой Microsoft Excel. Желательно также иметь хотя бы самое элементарное представление об основах теории статистики.

В процессе работы над книгой читатель сможет расширить свои знания о функциональных возможностях программы Excel, а также научиться работать в эконометрической программе EViews. Вполне естественно, что в первую очередь мы будем обращать внимание на обучение тем

функциям этих программ, которые потребуются для прогнозирования курса валют.

Начинающим пользователям, которые еще недостаточно опытны в работе с Microsoft Excel, можно порекомендовать прочитать книгу Н.И. Макарова, В.Я. Трофимец Статистика в Excel: Учеб. пособие. – М.: Финансы и статистика, 2003, а также другие учебные пособия (их довольно много), рассказывающие о возможностях данной программы. В свою очередь, тем читателям, которые хотят научиться прогнозировать курсы валют с помощью эконометрической программы EViews, можно порекомендовать в качестве пособия следующие книги: Молчанов И.Н., Герасимова И.А. Компьютерный практикум по начальному курсу эконометрики (реализация на EViews): Практикум /Рост. гос. экон. унив. – Ростов-на-Дону, – 2001, и Турунцева М.Ю. Анализ временных рядов / МИЭФ ГУ-ВШЭ. – М., 2003.

В заключении хочу сказать еще об одном важном условии, без которого работа с этой книгой будет не очень продуктивной: наличии у читателя желания овладеть приемами прогнозирования курса валют, которые автор постарался изложить максимально доступным языком.

С теми моими читателями, кому для лучшего усвоения материала книги потребуются дополнительные консультации, разъяснения или помощь в решении заданий, автор готов пообщаться по скайпу. Узнать об условиях получения консультаций по книге, а также записаться на консуль-

тацию можно по следующему адресу электронной почты:
bryukov@bk.ru

Владимир Георгиевич Брюков, независимый аналитик

Об авторе



Брюков Владимир Георгиевич, независимый финансовый аналитик, с 2003 года занимается банковской журналистикой.

С 2005 года особое место в его публикациях занимают статистические методы анализа валютных и финансовых рынков. Теме валютного прогнозирования, в первую очередь, прогнозу по курсу доллара США, посвящены многие его статьи, опубликованные в журналах «Валютный спекулянт», «Инвестиционный банкинг» и в ряде других изданий. В этих публикациях обобщаются результаты проведенного автором исследования валютного рынка, предлагаются оптимальные методы прогнозирования по курсам валют с учетом последних достижения современной статистической науки. С 2009 по 2015 год В. Г. Брюков ежемесячно публиковал прогнозы на портале Банкир.ру прогнозы по курсам пятнадцати ведущих мировых валют на будущий месяц. Насколько точными при этом были прогнозы, наши читатели могут убедиться сами, посетив на этом сайте рубрику «Валютный рынок». Сотрудничество с этим известным и авторитетным порталом, а также большой интерес, проявленный читателями к книге «Как предсказать курс доллара. Эффективные методы прогнозирования с использованием Excel и EViews», стали для автора важным стимулом, способствовавшим написанию новой книги по валютному прогнозированию.

Глава 1.

Понятие о стационарном и нестационарном временном ряде, выявление нестационарности ряда графическим способом

1.1. Краткая характеристика стационарных и нестационарных случайных процессов

Внезапно изменяющиеся тренды на валютном рынке, на первый взгляд, носят настолько причудливый и непредсказуемый характер, что многие инвесторы убеждены, что делать какие-то прогнозы по поводу курса валют – дело абсолютно безнадежное. И действительно, если посмотреть, например, на динамику ежемесячного курса доллара США (как впрочем, и на динамику других свободно конвертируемых валют), то этот временной ряд нельзя назвать стационарным. Чтобы понять, к каким последствиям – с точки зрения прогнозирования курса американской валюты – ведет этот факт, нам придется немного углубиться в теорию стационарных и нестационарных случайных процессов.

Как известно, в статистической литературе принято выделять три вида случайных процессов: строго стационарные, слабо стационарные и нестационарные процессы.

Случайный процесс, образующий временной ряд $X_1, X_2, X_3 \dots X_t$ (буква X обозначает переменную, содержащую определенную рыночную информацию, например, по динамике курсов валют, а, цифры – $1, 2, 3 \dots t$ – моменты времени), называется строго стационарным (или, как еще говорят,

стационарным в узком смысле), если совместное распределение вероятностей всех переменных $X_1, X_2, X_3 \dots X_t$ точно такое же, как и для наблюдений $X_{1+T}, X_{2+T}, X_{3+T} \dots X_{t+T}$ (где $t=t_2-t_1$ – временной лаг). Иначе говоря, свойства строго стационарного временного ряда не меняются при изменении начала отсчета времени.

Однако в сфере экономики, в том числе в сфере финансовых и валютных рынков, строго стационарные процессы отсутствуют, а потому для нас гораздо больший интерес представляют так называемые, слабые стационарные процессы или стационарные процессы в широком смысле. Под слабым стационарным процессом понимается случайный процесс, у которого среднее и дисперсия – независимо от рассматриваемого периода времени – имеют постоянное значение, а автоковариация зависит только от длины лага между исследуемыми переменными.

Напомним нашим читателям, что среднее значение временного ряда можно найти по следующей формуле (1.1):

$$X_{cp} = \frac{\sum (X_1 + X_2 + \dots + X_t)}{n}$$

где n – количество наблюдений во временном ряде.

Дисперсия (мера разброса случайной величины, например, отклонения курса доллара от его среднего значения, или, как еще говорят, от его математического ожидания) временного ряда представляет собой средний квадрат отклонений переменной (случайной величины) от ее среднего значения.

Соответственно, дисперсия находится по следующей формуле (1.2):

$$D = \frac{\sum (X_t - X_{cp})^2}{n}$$

Хочу заметить, что в Excel дисперсию можно найти, пользуясь функцией ДИСПР (если исходные данные представляют собой генеральную совокупность) или функцией ДИСП (если данные представляют собой выборку).

Для оценки тесноты и направления связи между переменными одного временного ряда с определенным лагом используется автоковариация. В частности, автоковариация

между значениями X_t и X_{t-T} отделенными друг от друга интервалом в T единиц времени, называется автоковариацией с лагом (задержкой) T , которая находится по следующей формуле (1.3):

$$Cov(X_t, X_{t-T}) = \frac{\sum (X_t - X_{cp.})(X_{t-T} - X_{cp.})}{n}$$

Автоковариацию, согласно формуле (1.3), в Excel можно найти с помощью функции КОВАР, которая возвращает величину ковариации. Причем, последнюю называют автоковариацией в том случае, когда ее используют для оценки тесноты и направления связи между переменными одного временного ряда с определенным лагом. Например, с лагом минус один месяц. Доказано, что для независимых переменных X и Y ковариация всегда равна нулю, а для зависимых переменных она, как правило, отличается от нуля. Если лаг $t = 0$, то автоковариация равна дисперсии.

Если временной ряд, характеризующий динамику, например, курса валют, является слабо стационарным, то это означает отсутствие: во-первых, тренда; во-вторых, строго периодических колебаний; в-третьих, систематических измене-

ний дисперсии; в-четвертых, каких-либо иных систематических изменений во временном ряде. Таким образом, под стационарным процессом в слабом или в широком смысле понимается случайный процесс, у которого среднее и дисперсия – независимо от периода времени – имеют постоянное значение, а автоковариация зависит от длины лага между рассматриваемыми переменными.

Если временной ряд является нестационарным, то с точки зрения теории это предполагает, что он содержит не только случайную компоненту, но и тренд, а его средняя, дисперсия и автоковариация изменяются во времени. В связи с этим делать прогнозы по нестационарному временному ряду более затруднительно (особенно на длительный период, или в период каких-либо резких изменений в его динамике), чем по стационарному ряду.

1.2. Распознавание стационарности временного ряда с помощью построения его графика

Существуют различные методы распознавания стационарности временного ряда, однако, пожалуй, самым простым из них является построение графика временного ряда с последующим визуальным определением наличия в нем тренда.

С этой целью мы решили построить график ежемесячных колебаний курса доллара к рублю за период с июня 1992 г. до апреля 2010 г. Читатели, которые не умеют строить диаграммы, могут ознакомиться с представленным ниже алгоритмом действий №1 и № 2

Алгоритм действий № 1 «Как строить диаграммы в Microsoft Excel»

Шаг 1. Поиск данных, их загрузка и первичная обработка в Excel

Во-первых, нужно на сайте Банка России <http://www.cbr.ru/> взять необходимые данные по ежедневным курсам доллара за весь интересующий нас период. Во-вторых,

после того как мы скопируем рыночную статистику в файл Microsoft Excel, все данные по курсу доллара к рублю с 1 июля 1992 г. по 1 января 1998 г. необходимо разделить на 1000, поскольку на сайте Банка России за этот период они приводятся в неденоминированном виде. В-третьих, для того, чтобы из всего массива данных оставить только необходимые для нас данные, а именно: курс доллара на конец месяца, необходимо их отфильтровать с помощью опций ДАННЫЕ/дополнительно/расширенный фильтр.

Шаг 2. Построение графика в Excel

Выделим с помощью мышки столбец с ежемесячными данными (на конец месяца) по курсу пары рубль – доллар за период с июня 1992 г. (на конец июня из-за отсутствия на сайте Банка России более ранних данных возьмем курс доллара на 1 июля 1992 г.) по апрель 2010 г. и столбец с соответствующими обозначениями месяцев. Далее выбираем в панели инструментов кнопку Вставка (в Excel 2007 г.), либо кнопку Мастер диаграмм (в Excel 1997-2003 гг.), в которой выбираем опцию График (см. рис.1.1).

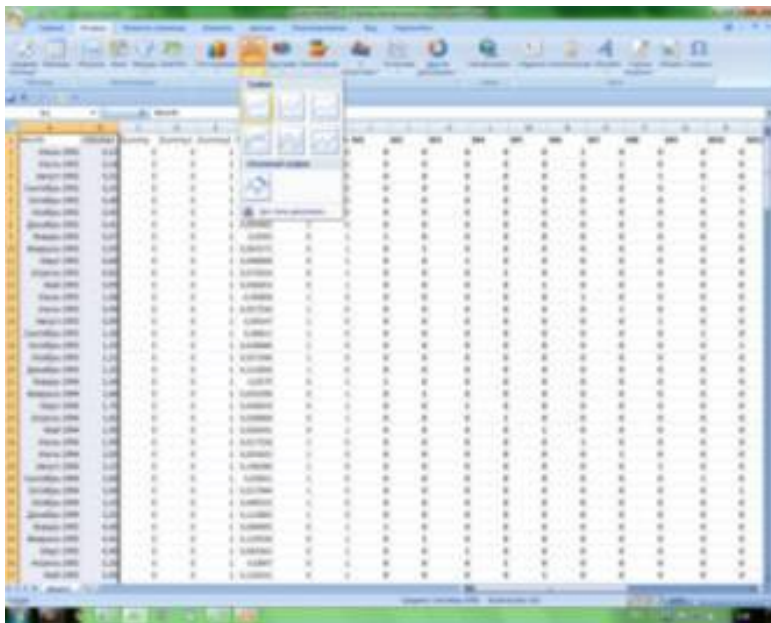


Рис. 1.1. Опция «График» – ВСТАВКА/МАСТЕР ДИАГРАММ

В результате у нас получился график (см. рис. 1.2), свидетельствующий о том, что динамику колебаний ежемесячного курса доллара нельзя назвать стационарной. Судя по данному графику, можно прийти к выводу, что во временном ряде наблюдается тенденция к росту, а среднее значение кур-

са доллара в разные периоды времени принимает различные значения. В частности, на графике хорошо видно, что во второй половине 1992 г. курс доллара, хотя и систематически рос, но в целом был лишь немного выше нулевой отметки. В то время как к концу 1998 г. он превысил уровень в размере 20 руб., а в 1999 – 2010 гг. курс американской валюты колебался в пределах от 24 руб. до 35 руб.



Рис. 1.2. Ежемесячный курс доллара США, в руб.

Источник: здесь и далее (если особо не оговорено) даются данные Банка России и расчеты автора

Теперь построим аналогичный график в EViews. Однако прежде нам нужно научиться импортировать данные в эту программу из исходного экселевского файла. Умение выполнять эту процедуру потребуется для последующей работы в EViews. С этой целью следует ознакомиться с алгоритмом действий № 2.

Алгоритм действий № 2 «Импорт данных и создание рабочего файла в EViews»

Шаг 1. Подготовка данных в Excel для их последующего импорта в EViews

Прежде чем приступить к созданию диаграммы в EViews нужно сначала импортировать в эту программу из Excel ежемесячные данные по курсу доллара к рублю. При работе в более ранних версиях EViews импортируемые данные необходимо сохранять в формате Excel 5.0/95 поскольку при использовании других экселевских форматов в EViews появится сообщение об ошибке. Однако в последних версиях EViews можно загружать данные из экселевских файлов в любом формате, в том числе и в форматах Excel 2007 и 2010 года.

Причем, импортируемые данные следует размещать в виде столбца в самой верхней строке экселевского листа слева. Например, заголовок первого столбца с данными должен

быть помещен в ячейке B1, заголовок второго столбца с данными – в ячейке C1 и так далее, в то время как заголовок с соответствующими датами – в ячейке A1.

Заголовки столбцов следует обозначать латинскими буквами, поскольку англоязычная программа EViews не понимает кириллицу. В частности, столбец с ежемесячными данными по курсу доллара США мы решили обозначить как USDOLLAR (поместили в ячейку B1), а заголовок (в ячейке A1) с названиями месяцев – Month. И последний важный момент: экселевский лист, на котором мы разместим подготовленные к импорту в EViews данные, нужно также переименовать латинскими буквами. В данном случае мы назвали экселевский лист с импортируемыми данными sheet1.

Шаг 2. Создание рабочего файла в EViews

Для того чтобы создать рабочий файл, содержащий данные, с которыми мы собираемся работать, необходимо в главном меню EViews выбрать опции File/New/Workfile. В результате откроется следующее диалоговое мини-окно (см. рис. 1.3.):

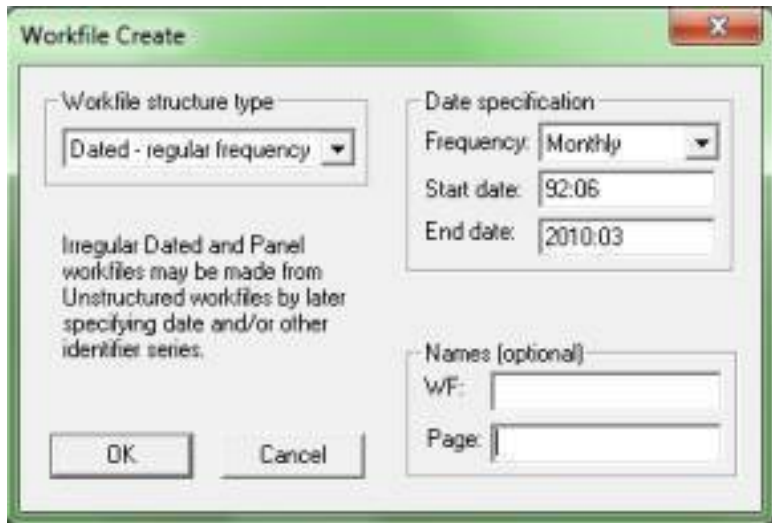


Рис. 1.3. Диалоговое окно Workfile create

В этом диалоговом окне необходимо задать соответствующую информацию. Так, в мини-окне Workfile structure type (структура рабочего файла) мы задаем опцию Dated-regular frequency (даты с определенной частотой). Соответственно, в мини-окне Frequency (частота данных) ставим опцию Monthly (ежемесячные данные), в Start date (начальная дата) – 92:06 (июнь 1992 г.), в End date (конечная дата) – 2010:03 (март 2010 г.). Хочу заметить, что в мини-окне End date дату года нужно обязательно давать четырехзначной, в то вре-

мя как в Start date она может быть двузначной. В результате у нас получится (см. рис.1. 4) неполный рабочий файл (Workfile): в нем будут отсутствовать данные, которые еще предстоит импортировать.

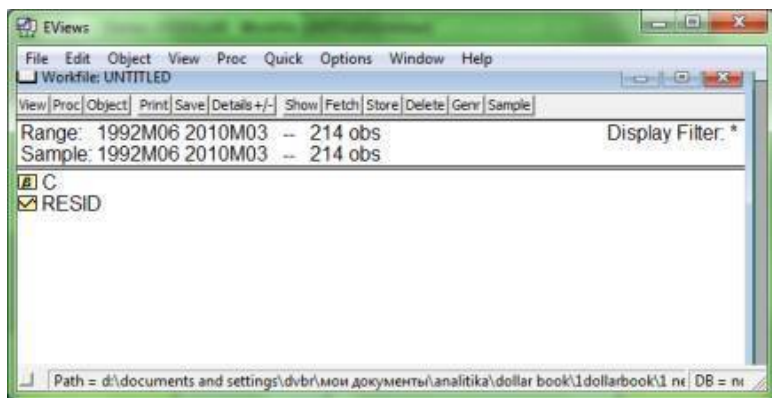


Рис. 1.4. Неполный рабочий файл Workfile

Шаг 3. Импорт данных в EViews

Перед импортом данных экселевский файл нужно обязательно закрыть, поскольку иначе появится сообщение об ошибке. При работе в последних версиях EViews в командной строке этой программы нужно воспользоваться опции-

ей IMPORT/IMPORT FROM FILE. После открытия экселевского файла появится окошко Excel read-Step 1 of 3 (см. рис. 1.5), в котором следует выбрать одну из двух опций: Predefined range (предопределенный диапазон) или Custom range (обычный диапазон). В случае выбора Predefined range в EViews загружаются уже выбранные программой данные, а если вы воспользуетесь опцией Custom range, то в этом случае можно самому выбрать требуемый диапазон данных, в том числе внести необходимые правки в опции SHEET (лист), START CELL (начальная ячейка), END CELL (конечная ячейка).



Рис. 1.5. Окно Excel 97-2003 Read – Step 1 of 3

Если вы работаете в более ранних версиях EViews, то при импорте данных в командной строке нужно выбрать опции File/Import/Read Text-Lotus-Excel. После этого появится новое диалоговое окно Excel Spreadsheet Import (импорт листа Excel). В открывшемся диалоговом окне (см. рис. 1.6) нужно отметить в мини-окне Excel5+sheet name – название листа, которое у нас обозначено как sheet1. В мини-окне Names for

series or Number if named in file (название для серии данных или номер серии данных, если у нее есть название в файле) – поставим цифру 1, поскольку мы импортируем лишь одну серию данных, которую уже называли USDOLLAR. В остальных мини-окнах соответствующие опции в EViews устанавливаются по умолчанию. В частности, в мини-окне UPPER-LEFT DATA CELL (верхняя левая ячейка с данными) – по умолчанию указывается ячейка B2.

Более подробно обо всех нюансах импорта данных из Excel в EViews можно прочитать, например, в книге Турунцевой М.Ю. Анализ временных рядов / МИЭФ ГУ-ВШЭ. – М., 2003 стр. 4-9.



Рис. 1.6. Диалоговое окно Excel Spreadsheet Import (импорт листа Excel)

Поскольку мы уже создали рабочий файл в EViews, то построить график курса доллара, аналогичный экселевскому (см. рис. 1.2), не представляет особого труда. В рабочем файле (Workfile) EViews открываем файл USDOLLAR, после чего используем для построения диаграммы в виде графика (LINE) опции VIEW/GRAPH/LINE (см. рис. 1.7).

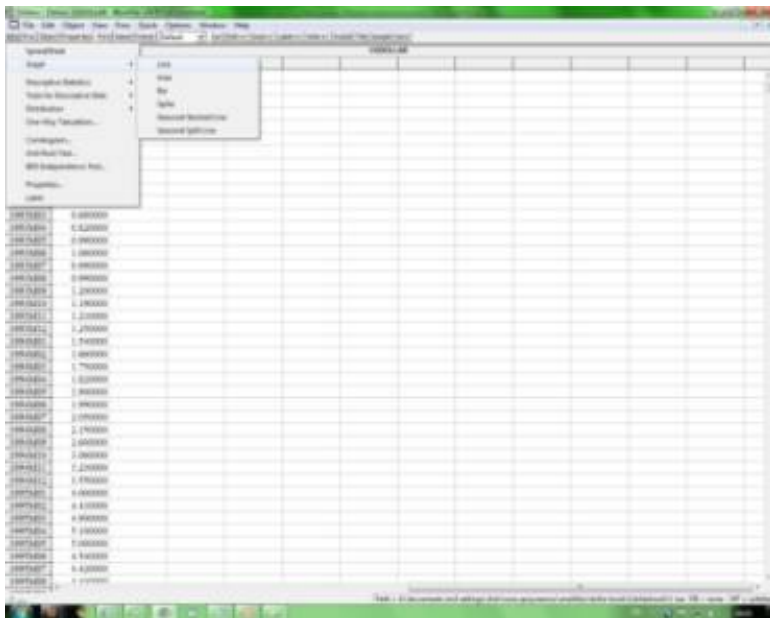


Рис. 1.7. Использование опции VIEW/GRAPH/LINE для построения в EViews линейного графика LINE

В результате у нас получилась диаграмма (рис. 1.8), вполне аналогичная (если не считать различный тип форматирования) диаграмме на рис 1.2, построенной в Excel. Для того чтобы сохранить полученную диаграмму в EViews на отдель-

ном листе следует нажать верхнюю кнопку FREEZE (окончательно принять).

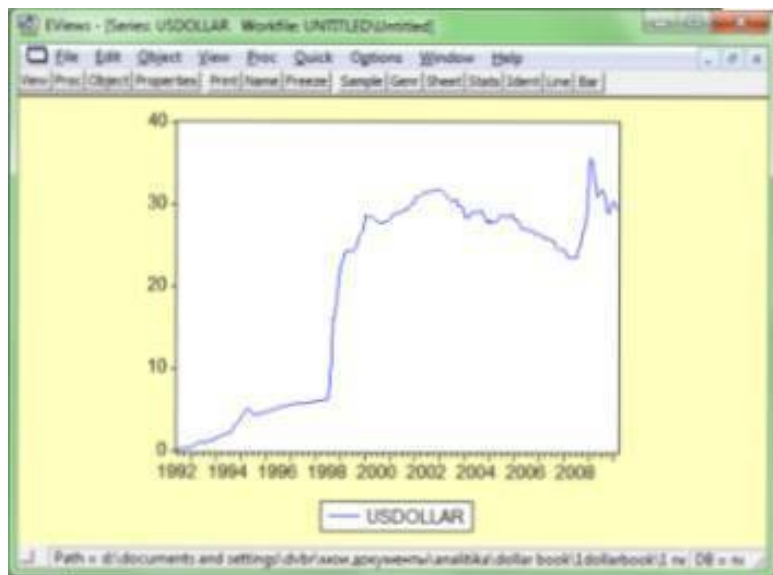


Рис 1.8. График курса доллара, полученный в EViews

Таким образом, построив соответствующие графики в EViews и Excel, нам удалось выяснить, что временной ряд, характеризующий динамику ежемесячного курса доллара,

является нестационарным, поскольку в нем наблюдается ярко выраженный тренд. Вместе с тем, как мы уже говорили ранее, нестационарный временной ряд содержит не только тренд, но и случайную компоненту. Следовательно, чтобы сделать адекватный прогноз по курсу доллара необходимо учесть как тренд, так и случайную компоненту, поскольку оба эти фактора существенно влияют на динамику валюты.

Схематично, наша дальнейшая работа, которой посвящены последующие главы этой книги, будет заключаться в следующем. Во-первых, нам нужно составить уравнение регрессии, с помощью которого можно будет делать прогнозы с необходимой точностью. Во-вторых, необходимо протестировать данное уравнение регрессии (прогностическую модель) на его адекватность с точки зрения прогностических качеств. В-третьих, надо составить точечные прогнозы по курсу американской валюты, используя полученную математическую модель. В-четвертых, нужно удостовериться в приемлемой точности составленных точечных прогнозов. В-пятых, необходимо убедиться, что получившиеся в результате отклонения фактического курса доллара от его предсказанных (расчетных) значений представляют собой стационарный ряд. В-шестых, надо посмотреть, является ли распределения остатков нормальным, что позволит нам впоследствии составить интервальные прогнозы – с учетом диапазона отклонений точечных прогнозов от фактического курса доллара – с определенным уровнем надежности. В-седьмых, нуж-

но проверить, соответствует ли точность интервальных прогнозов заданному уровню надежности. В-восьмых, научиться применять полученную статистическую модель для составления рекомендуемых цен покупки и продажи валюты, используемых в качестве стоп-приказов при работе на валютном рынке. При этом выполнение всех этих процедур будет сопровождаться подробным рассказом о том, как их можно выполнить в Excel или EViews, что поможет нашим читателям впоследствии самостоятельно решать эти задачи.

Контрольные вопросы и задания к главе 1

1. Чем отличаются строго стационарные процессы от стационарных процессов в широком смысле?
2. Может ли стационарный процесс иметь тренд или какие-либо строго периодические колебания?
3. Чем нестационарный процесс отличается от стационарного? Может ли у нестационарного процесса быть тренд?
4. Если Вы пришли к выводу о нестационарности данного временного ряда, то, что можно сказать об устойчивости его средней, дисперсии и автоковариации. Дайте определение средней, дисперсии и автоковариации.

Глава 2.

Метод наименьших квадратов и решение уравнения регрессии в Excel

2.1. Характеристика метода наименьших квадратов и его применение при прогнозировании курса доллара

Как мы выяснили в главе 1, динамика курса валют представляет собой временной ряд, имеющий не только тренд, но и случайную компоненту, поэтому в качестве метода оценки параметров прогностической модели, как правило, используется регрессионный анализ. Как известно, задачей регрессионного анализа является определение аналитического выражения (математической формулы), аппроксимирующего связь между зависимой переменной Y (ее называют также результативным признаком) и независимыми (их называют также факторными) переменными $X_1, X_2 \dots X_n$. При этом форма связи результативного признака Y с факторами $X_1, X_2 \dots X_n$, либо с одним фактором X , получила название уравнения регрессии. В качестве метода аппроксимации (приближения) в уравнении регрессии используется метод наименьших квадратов (МНК), который минимизирует сумму квадратов отклонений фактических значений Y от его предсказываемых значений, рассчитанных по опреде-

ленной математической формуле. Причем, решение уравнения регрессии относительно интересующих нас переменных y (курс доллара) и x (время или порядковый номер месяца), по сути, заключается в подборе прямой линии к совокупности данных, состоящих из пар данных, характеризующих динамику курса доллара и соответствующие порядковые номера месяцев. При этом линию, которая лучше всего подойдет к этим данным, выбирают так, чтобы сумма квадратов значений вертикальных отклонений зависимой переменной (фактического курса доллара) от линии, рассчитанной по уравнению регрессии (предсказанный курс доллара), была минимальной.

Математические подробности оценки параметров уравнения регрессии методом наименьших квадратов

В самом общем виде формулу МНК можно представить следующим образом (2.1):

$$\sum (Y_t - Y_{расч.})^2 \Rightarrow \min. \sum e^2$$

где Y_t и $Y_{\text{расч.}}$ – фактические и расчетные значения зависимой (результативной) переменной для различных моментов времени;

$$\min. \sum e^2$$

– минимальная сумма квадратов отклонений (остатков) фактических значений Y от его расчетных (предсказываемых) значений.

Поскольку $Y_{\text{расч.}} = a + bX$ (где a – свободный член уравнения регрессии, а b – коэффициент регрессии), то уравнение (2.1) примет следующий вид (2.1.1):

$$\sum (Y_i - a - bX)^2 \Rightarrow \min. \sum e^2$$

Для отыскания параметров a и b , при которых функция $f(a,b)$ принимает минимальное значение, необходимо найти частные производные по каждому из параметров этой функ-

ции a и b и приравнять их нулю. Если минимальную сумму квадратов отклонений (остатков) e^2 обозначить через S , то в результате мы получим систему нормальных уравнений МНК для прямой (2.1.2):

$$\frac{dS}{da} = -2 \sum Y + 2na + 2b \sum X = 0$$

$$\frac{dS}{db} = -2 \sum YX + 2a \sum X + 2b \sum X^2 = 0$$

Преобразовав систему уравнений (2.1.2) получим (2.1.3):

$$na + b \sum X = \sum Y$$

$$a \sum X + b \sum X^2 = \sum XY$$

Решив систему уравнений (2.1.3) методом последователь-

ного исключения переменных найдем следующие оценки параметров:

$$a = \bar{Y} - b\bar{X}; \quad (2.1.4)$$

$$b = \frac{\overline{YX} - \bar{Y} \times \bar{X}}{(X^2)_{\text{ср}} - \bar{X}^2}; \quad (2.1.5)$$

где \bar{Y} , \bar{X} , $(X^2)_{\text{ср}}$ и \overline{YX} — средние значения переменных X , Y , X^2 и их произведения XY .

С помощью оцененного таким образом уравнения регрессии можно предсказать, как в среднем изменится признак Y в результате роста факторов $X_1, X_2 \dots X_t$ (или одного фактора X). В зависимости от того, какая математическая функция используется для прогнозирования результирующей переменной Y , различают линейную и нелинейную регрессию. При этом в основе линейной регрессии лежит уравнение линейного тренда, а в основе нелинейной регрессии – целое семейство уравнений нелинейных трендов (полиномиальный второй, третьей и прочих степеней, степенной, экспоненциальный, логарифмический и другие). В случае если результирующий признак Y зависит от одного фактора X , то такое уравнение регрессии называется парным, а если Y зависит от нескольких факторов $X_1, X_2 \dots X_t$ – то уравнением множественной регрессии.

Практически в любом учебнике по общей теории статистики и по эконометрике можно более подробно познакомиться со спецификой уравнений регрессии. (См., например, учебник «Эконометрика» под ред. И.И. Елисеевой. – 2-е изд., пер. и доп. – М.: Финансы и статистика, 2006, стр. 43-132).

Существуют формулы, по которым можно самостоятельно найти параметры, как уравнения линейной регрессии, так и различных видов уравнений нелинейной регрессии. Однако с внедрением в широкую практику компьютеров и соответствующих компьютерных программ уже нет необходимости оценивать параметры уравнения регрессии вручную, тем более что этот процесс довольно трудоемкий.

2.2. Решение уравнения регрессии в Excel с учетом фактора времени. Интерпретация и оценка значимости полученных параметров

Поэтому далее остановимся на изучении алгоритма решения уравнений регрессии с применением соответствующих вычислительных программ. При этом работу с уравнениями регрессии в компьютерных программах можно разделить на три этапа.

На первом, подготовительном этапе необходимо определиться с набором факторов, которые необходимо включить в уравнение регрессии, а также с его аналитической формой, что в ряде случаев требует предварительной обработки данных. Например, в случае выбора степенного уравнения регрессии вместо исходных данных нужно взять их логарифмы.

Второй этап состоит из собственно решения уравнения регрессии и нахождения его параметров.

На третьем этапе проводится оценка и тестирования общего качества уравнения регрессии, проверка статистической значимости каждого из коэффициентов регрессии,

определяются их доверительные интервалы, а также принимается окончательное решение об адекватности или неадекватности полученного уравнения регрессии.

Как известно, одним из наиболее распространенных способов определения тренда в динамике курса валюты является построение его зависимости от фактора времени T . Так, если в качестве зависимой переменной Y мы возьмем ежемесячный курс доллара, а в качестве независимой переменной T – время (в данном случае порядковые номера месяцев, начиная с июня 1992 г.=1), то у нас получится следующее уравнение парной линейной регрессии:

$$Y_{\text{расч.}} = a + bT \quad (2.2);$$

где a – свободный член уравнения регрессии; b – линейной коэффициент регрессии, показывающий, как изменение величины независимой переменной (фактора) T в среднем способствует изменению зависимой переменной (результативного признака) Y ; $Y_{\text{расч.}}$ – расчетное значение результативного признака, вычисляемое по формуле (2.2).

Минимизируем сумму квадратов отклонений (остатков) $Y_{\text{факт.}}$ от $Y_{\text{расч.}}$, то есть от фактических значений курса доллара от его расчетных значений. В результате формулу МНК (2.1.1) для линейной регрессии можно в данном случае представить в виде формулы (2.3):

$$\sum [Y_{\text{факт}} - (a + bT)^2] \Rightarrow \min \sum e^2$$

Уравнение (2.3) в принципе можно решить самостоятельно, если найти его параметры согласно формулам (2.1.4) и (2.1.5), но в целях ускорения этого процесса мы будем его решать с помощью Пакета анализа Excel. Кстати, желающие лучше усвоить суть МНК могут сначала самостоятельно в «ручном режиме» решить данное уравнение регрессии, а затем сверить свои результаты с теми, что мы получим в Excel.

Для того чтобы подготовить исходные данные к решению данного уравнения регрессии разместим в Excel два столбца исходных данных. В первом столбце, который озаглавим Time, поместим порядковые номера месяцев, начиная с июня 1992 г. (с номером =1) и кончая апрелем 2010 г. (с номером =215). Во втором столбце, который озаглавим USDOLLAR, поместим данные по курсу доллара на конец месяца, начиная с июня 1992 г. и заканчивая апрелем 2010 г. (последние данные, имевшиеся на тот момент, когда писались эти строки). Таким образом столбец Time представляет собой независимую переменную, которая в формуле (5) обозначена символом T, а столбец USDOLLAR является зависимой переменной Y_{факт}. Далее переходим к решению уравнения регрессии в Пакете анализа Excel, о том, как это

делается, можно прочитать ниже – в алгоритме действий № 3.

Алгоритм действий № 3 «Как решить уравнение регрессии в Excel»

Шаг 1. Ввод в уравнение исходных данных

Делается это следующим образом: сначала в Microsoft Excel 2007 г. в верхней панели инструментов выбирается опция Данные (в Microsoft Excel 1997-2003 гг. нужно выбрать опцию Сервис), потом в появившемся окне Анализ данных – опция Регрессия. После чего появляется новое окно – Регрессия (см. рис. 2.1), в котором в графе Входной интервал у выделяем (с помощью мышки) столбец данных USDOLLAR (ячейки $\$C\$1: \$C\216). Здесь же в графе Входной интервал X» выделяем столбец данных Time(ячейки $\$B\$1: \$B\216), то есть независимую переменную T из нашего уравнения регрессии (5).

Шаг 2. Дополнительные опции

Если бы мы хотели получить уравнение регрессии без свободного члена, который в формуле (2.2) обозначен символом а, то тогда нам следовало бы выбрать еще и опцию КОНСТАНТА-НОЛЬ. Однако в данном случае в использовании этой опции нет необходимости.

Опцию Остатки следует выбирать тогда, когда есть необ-

ходимость, чтобы в выходных данных содержалась информация об отклонении расчетных y от их фактических значений. При этом остатки находятся по следующей формуле (2.4):

Остатки = $Y_{\text{расч.}} - Y_{\text{факт.}}$ (7); где $Y_{\text{расч.}}$ – расчетные, $Y_{\text{факт.}}$ – фактические значения результативного признака.

Опцию МЕТКИ применяют для того, чтобы переменные, включенные в уравнение регрессии, в выводе итогов были обозначены в виде заголовков соответствующих столбцов.

По умолчанию оценка в Excel параметров уравнения регрессии делается с 95% уровнем надежности. Однако в случае необходимости в опции Уровень надежности можно поставить цифру 99, что означает задание для программы оценить коэффициенты регрессии с 99% уровнем надежности. В результате в выводе итогов мы получим данные, характеризующие как в целом уравнение регрессии, так и верхние и нижние интервальные оценки коэффициентов данного уравнения с 95% и 99 % уровнями надежности. При 95% уровне надежности существует риск, что в 5 % случаях оценки коэффициентов уравнения регрессии могут оказаться неточными, а при 99% уровне надежности этот риск равен 1%.

Шаг 3. Вывод итогов

На заключительном этапе выбираем в параметрах выво-

да (окно РЕГРЕССИЯ) опцию выходной интервал, в которой указываем соответствующую ячейку Excel (\$H\$2), далее щелкаем по надписи ОК и получаем ВЫВОД ИТОГОВ (см. рис 2.1, где можно увидеть все заданные нами параметры уравнения регрессии). В случае необходимости вывод итогов можно получить на отдельном листе (см. опцию НОВЫЙ РАБОЧИЙ ЛИСТ) или в новой книге Excel (см. опцию НОВАЯ РАБОЧАЯ КНИГА).

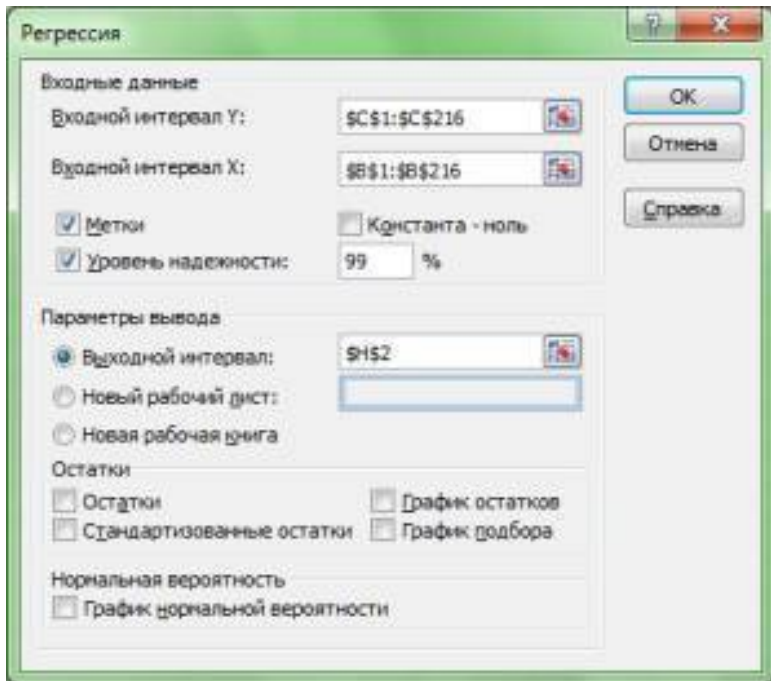


Рис. 2.1. Диалоговое окно РЕГРЕССИЯ для вывода итогов при решении в Excel уравнения регрессии

Результаты решения уравнения регрессии, которые в программе Excel выдаются в виде единой таблицы под заголовком **ВЫВОД ИТОГОВ**, у нас представлены в виде трех блоков (см. табл. 2.2-2.4). Так, в табл. 2.2 сгенерированы резуль-

таты по регрессионной статистике, в табл. 2.3 дается дисперсионный анализ, а в табл. 2.4 оценивается статистическая значимость коэффициентов регрессии.

Параметры, представленные в табл. 2.2, оценивают уровень аппроксимации фактических данных, полученный с помощью данного уравнения регрессии. Так, параметр Множественный R обозначает множественный коэффициент корреляции R, который характеризует тесноту связи между результативным признаком Y и факторами переменными X1, X2...Xn. Данный коэффициент изменяется в пределах от 0 до 1, причем, чем ближе к 1, тем теснее корреляционная связь между переменными, включенными в уравнение регрессии. Множественный коэффициент корреляции равен квадратному корню, извлеченному из коэффициента детерминации R², который у нас также приводится в регрессионной статистике. Множественный коэффициент R также находят по формуле (2.5):

$$R = \sqrt{\frac{\sum (Y_{\text{расч}} - Y_{\text{ср.}})^2}{\sum (Y_{\text{факт}} - Y_{\text{ср.}})^2}}$$

где Y факт. – фактическое, а Y расч. – расчетное (пред-

сказанное по уравнению регрессии) значение результативного признака.

Зная величину коэффициента корреляции R , можно дать качественную оценку силы связи между зависимой и независимыми переменными, включенными в данное уравнение. С целью классификации силы связи обычно используют шкалу Чеддока (см. табл. 2.1).

Таблица 2.1. Шкала Чеддока для классификации силы связи

Величина коэффициента множественной корреляции R	Оценка силы связи
0,1—0,3	Слабая
0,3—0,5	Умеренная
0,5—0,7	Заметная
0,7—0,9	Высокая
0,9—0,99	Весьма высокая

В случае между переменными существует функциональная связь, то $R=1$, а если корреляционная связь отсутствует, то $R=0$. Поскольку в таблице 2.2 множественный коэффициент корреляции R равен 0,8456, то согласно таблице Чеддока, связь между переменными, включенными в уравнение регрессии можно считать высокой. Следует также заметить,

что если коэффициент множественной корреляции меньше 0,7, то это означает, что величина коэффициента детерминации R^2 (о нем мы расскажем ниже) будет меньше 50%, а потому регрессионные модели с таким коэффициентом детерминации не имеют большого практического значения.

Однако самым важным является другой параметр регрессионной статистики – **R-квадрат** (его мы выделили жирным шрифтом), обозначающий коэффициент детерминации R^2 . Коэффициент детерминации R^2 характеризует долю дисперсии результативного признака Y , объясняемую уравнением регрессии, в общей дисперсии результативного признака. Коэффициент детерминации R^2 находится по формуле (2.6):

Конец ознакомительного фрагмента.

Текст предоставлен ООО «ЛитРес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на ЛитРес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.